



Victor Chang
Cardiac Research Institute

Database for Cardiovascular Functional Genomics

Adam Hill

ARDC Data and Services summit / October 2019



Dr Victor Chang AC 1936-1991, Pioneering Cardiothoracic Surgeon and Humanitarian

The project



- Engagement program/workshop to scope a data resource bringing together genomic and phenotypic data from cardiovascular research labs across Australia.
- Specific focus on functional data (electrophysiology, calcium handling, metabolism etc) from iPSC cardiomyocytes to answer questions about disease processes, population variability in drug responses, and identification of new therapies.
- Workshop involving national leaders from relevant disciplines was held on 27th August: ***'Future-Proofing Cardiovascular Dataset Management: Identifying and Implementing a Collaborative National Approach'***
- Organisations involved:
 - Australian Cardiovascular Alliance (ACvA);
 - Australian eResearch Organisation (AeRO);
 - Australian Functional Genomics Network;
 - NHMRC Accredited Academic Health Research Translation Centres;
 - Medical Research Institutes (VCCRI, Centenary Institute WIMR);
 - Local Health Districts (Sydney, Northern Sydney, Western Sydney).
 - Universities
- ~40 participants (NSW, QLD, VIC, SA, WA), peak bodies representing 100s researchers across Australia as well as international consultation

Key Issues

- 1) Australia has world leading expertise and resources in this domain but data is siloed with little interaction or coordination between groups. While this is currently a limitation, it also presents a significant opportunity.
- 2) FAIR data – Starting from a very low bar. Data is siloed, not available online, often proprietary, no metadata or licencing.
- 3) There is no standardisation of data formats or metadata in our domain - this was perhaps the main practical hurdle identified. It was proposed that a minimum information standard for phenotypic data should be published as part of the establishment of the collection.
- 4) Funding is another major limiting factor in establishing the data collection. Suitable sources identified were: NSW OHMR cardiovascular capacity building fund, MRFF Mission for cardiovascular health, the Ramsey Foundation, NHMRC and the ARDC platforms fund.
- 5) The Cardiovascular community (and biological community in general) has limited knowledge about what infrastructure (eg. NCRIS funded, cloud) is available. Education/outreach and/or facilitation of interactions with cloud platforms would provide a useful gateway.
- 6) Although on the whole there is broad acceptance of the advantages of collaborative contribution to a shared collection, there is still some community resistance in life sciences.

Lessons Learnt 1

- 1) Community requirement is not for raw data, but analysed datasets and/or concierged access.
 - Analysed data should be the focus of a central resource with standardised raw data linked and available to download.

- 2) The usefulness/integrity of the proposed data collection is dependent on standardisation of biological samples (not just data collection from the samples).
 - May require mandating of protocols for lab workflows – likely to be resistance.

- 3) Ethical requirements for patient-related samples and/or to enable linkage to other patient records are complex but enhance impact/usability.
 - Keep ethics as broad as possible, implement early, and learn from existing biobanks/linkage spines.

Lessons Learnt 2

-
- 4) Administrators of similar databases (eg. Procan: International Centre for the Proteome of Cancer Hosting) reported willingness of eg. AWS to assign resources and expertise to facilitate startup
→ Preference for hosting on cloud provider.
- 5) Timeliness – A window of opportunity exists around MRFF both in terms of funding and biological material/data
→ Coordinate with Australian Functional Genomics network, Australian Genomics Cardiovascular Flagship and AcVA Big data and Bioengineering flagships around researcher engagement and funding.
- 6) Specific short-term priority actions were identified as:
- i. Establishment and publishing of minimum data standard;
 - ii. Appointment of a governance team;
 - iii. Identification of funding opportunities;
 - iv. Engagement with providers (Cloud and/or NCRIS infrastructure) ;
 - v. Recruitment of admin/coordinators/developers;

Acknowledgements

Speakers

- Prof Louisa Jorm (Director, Centre for Big Data Research in Health, UNSW)
- Warran Kaplan (Chief of Informatics, Garvan Institute of Medical research)
- Tom Honeyman (ARDC)
- Luc Betbeder-Matibet (Director Research Technology Services, UNSW)
- Dr Tom Briffa (University of Western Australia)
- Dr David Elliot (Murdoch Children's Research Institute)
- Prof Sally Dunwoodie (Australian Functional Genomics Network)
- Prof Angus Lamond (hiPSCI)



Sam Moskwa

Loretta Davis

